"Made available under NASA sponsorship
In the interest of early and wide dissemination of Earth Resources Survey
Program information and without liability
for any use made thereof."

E78-10198
TM-79742

National Aeronautics and Space Administration

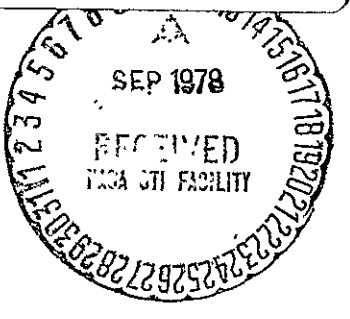
EARTH RESOURCES LABORATORY

PROCEDURES FOR GATHERING GROUND TRUTH INFORMATION FOR A SUPERVISED APPROACH TO A COMPUTER IMPLEMENTED LAND COVER CLASSIFICATION OF LANDSAT ACQUIRED MULTISPECTRAL SCANNER DATA

REPORT NO. 163

MARCH, 1977

(E78-10198) PROCEDURES FOR GATHERING GROUND TRUTH INFORMATION FOR A SUPERVISED APPROACH TO A COMPUTER IMPLEMENTED LAND COVER CLASSIFICATION OF LANDSAT ACQUIRED MULTISPECTRAL SCANNER DATA (NASA) 83 p
N78-31496
Unclas
G3/43 00198



LYNDON B. JOHNSON SPACE CENTER

The EARTH RESOURCES LABORATORY is an element of the LYNDON B. JOHNSON SPACE CENTER, HOUSTON, TEXAS: and is located at 1010 Gause Blvd., SLIDELL, LOUISIANA 70458. Telephone 504 255-6511 or FTS 685-6511.

"Made available under NASA sponsorship
In the interest of early and wide dissemination of Earth Resources Survey
Program information and without liability
for any use made thereof."

PROCEDURES FOR GATHERING GROUND TRUTH INFORMATION FOR A SUPERVISED APPROACH TO A COMPUTER IMPLEMENTED LAND COVER CLASSIFICATION OF LANDSAT ACQUIRED MULTISPECTRAL SCANNER DATA

By

Armond T. Joyce

March, 1977

TABLE OF CONTENTS

I. Introduction..... 1

II. Background and Basic Theory..... 2

III. Pre-Field Implementation Procedures..... 27

IV. Field Implementation Procedures..... 38

V. Concluding Comments..... 57

Appendix A Land Cover Conditions in Mississippi..... 61

B Definitions and Criteria..... 63

C Ground Truth Forms..... 66

D Instruction Sheet..... 75

E Verification Pattern..... 77

F Revisit Form..... 79

I. INTRODUCTION

This report is one of a series of reports that address a project conducted by the NASA Earth Resources Laboratory and the State of Mississippi Office of Science and Technology in cooperation with other State of Mississippi agencies. The overall project is entitled "Natural Resources Inventory System ASVT" (Application Systems Verification and Transfer), and has two facets. One facet involves the transfer of technology associated with the use of Landsat (formerly Earth Resources Technology Satellite) digital data and computer-implemented techniques for resource inventory. The other facet encompasses the demonstration of various specific applications for which the system has utility. In the series of reports that address the project, separate reports will cover the demonstrated applications in agriculture, forestry, wildlife management, etc. In addition, other separate reports in the series will cover the hardware description, software documentation, and procedures utilized during the implementation of the system. This particular report addresses the procedures used for ground truth information gathering. The report will outline the procedures used in the course of this project, but an attempt will also be made to address alternative procedures as well as basic theory so as to provide information to others who are planning ground truth information gathering operations for the purpose of using Landsat digital data for resource inventories. Although it is not necessary for personnel engaged in ground truth information gathering to

understand all ramifications of using Landsat digital data for resource inventory, the quality of their work will be enhanced by an understanding of the basic principles involved. Consequently, an attempt will be made to draw a line between what is considered "basic," and therefore covered in this report, and what is not "basic." However, because it is often difficult to draw such lines, this report will frequently cite other literature to which the reader may wish to refer for additional details.

II. BACKGROUND AND BASIC THEORY

After the acquisition of computer-compatible tapes (CCT's) that contain the raw data acquired by the multispectral scanner on the Landsat satellite, the first step in data processing during the project involved the use of a module of six computer programs developed at the Earth Resources Laboratory (ERL) and named PATREC (Pattern Recognition Analysis). The basic function of the PATREC computer programs is to effect a computer-implemented classification of each "cell" which represents 1.1 acres on the Earth's surface, for which data have been acquired by the multispectral scanner on the Landsat satellite.[1] This "classification" results in each of these 1.1 acre areas being categorized as pertaining to some land cover category; e.g., pine forest,

[1]A data "cell" is also referred to as a "pixel," a "data element," or a "resolution cell" in other literature, and relates to the instantaneous field-of-view of the multispectral scanner.

soybean, sand beach, etc., that the computer was trained to recognize.

The computer programs comprising ERL's PATREC module relate to the "supervised" technique, and the classifier algorithm is based on maximum likelihood ratio calculation and Bayesian decision rules. (See Whitley, S. L., 1975 and Jones, C., 1974 for additional theory and details.) The supervised technique requires that the location of a number of sites on which the land cover is known (e.g., a soybean field) be established in the Landsat data. These areas which are selected to contain a uniform, homogeneous land cover (e.g., a soybean field that is uniform in respect to planting date, density, vigor, etc.) are called "training sample sites," because, in a simplistic sense, they are used as references to "train" the computer to recognize the same land cover elsewhere. It is the office and field activities associated with establishing the true ground cover composition of these "training sample sites" that are encompassed by the ground truth information gathering operation.

Training Sample Site Criteria

Training sample site criteria must be established in recognition of how vegetation/land cover variables influence the energy being measured by the multispectral scanner (MSS) as energy is reflected or emitted from the land surface.

There are many characteristics of plants that determine the intensity and wavelength of reflected energy. Some of

these are size of plant cells, thickness of cell walls intercellular air spaces, leaf arrangement on the stem, pigments present, leaf water content, leaf pubescence, thickness and shape of leaf, etc. Inasmuch as a given plant species is usually unique in respect to these characteristics, the plant species is one variable which must be addressed. Some vegetation types such as agricultural crops, planted grasses, some forest plantations, and orchards are likely to consist of a single species; however, naturally occurring vegetation is usually a mixture of various species. Consequently, natural vegetation cover types are defined and named in respect to the predominant species. For example, a forest may be called a pine forest if 75% or more of the surface area of the tree crowns in the upper canopy were pine trees. Therefore, a training sample site selected to represent a pine forest cover type, as defined in this example, could include some hardwood trees if they were uniformly intermingled with the pine trees and their crowns did not cover more than 25% of the total surface area. It is also possible to define and name a vegetation cover type in respect to two or three species that grow in association with one another but together are predominant. For example, a forest cover type may be called oak-hickory if oak and hickory grew together in an intermingled manner and together comprised 75% or more of the surface area covered by tree crowns. Again, as with the pine forest example, a training sample site established to represent

the oak-hickory cover type could include other species but should meet the criteria and be uniform in respect to how the various species were intermingled. Although forest vegetation was referred to in the previous examples, the same statements concerning vegetation type criteria apply to marsh (non-forested wetlands) vegetation and brushland (multi-stemmed, woody shrubs) vegetation.

Although a given plant species is likely to be uniquely different from another plant species in respect to the various plant characteristics mentioned earlier, it is possible that some of those characteristics can change with plant age or plant vigor. For example, a young, vigorously growing plant may have less leaf water content or may have cells with thinner cell walls than an older, slow-growing plant of the same species. Consequently, the plant's age/vigor is the second variable that must be addressed, especially in the case of perennial vegetation. Of these two parameters, the vigor, as indicated by the rate at which the plant is growing, is the more important. It is referred to in conjunction with age only because there is a general correlation between vigor and age. In forested areas, there is likely to be a gradient from very young, vigorously growing forest stands to mature, stagnant, or even decadent forest stands as well as some all-aged stands. It is important that training sample sites be established to encompass all age/vigor variation in each vegetation type to be addressed; however, for practical purposes, it is

recommended that the age/vigor categories be fairly broad. For example, such categories may include (1) those young stands that are on good sites in respect to soil and rainfall and, therefore, are growing at a fast rate usually characterized by the profuse flushes of terminal branch growth during the spring, (2) those stands that have a moderate rate of growth because of site conditions and/or age, and (3) those stands in which growth has slowed down appreciably because the trees are near maturity, mature, stagnated, or on poor site conditions. As an example, in the case of Mississippi pine forests, the first category is likely to include plantations that are generally in the 1 to 10 year age bracket and are not yet of commercial size; the second category would include mainly pulpwood size trees generally in the 11 to 30 year age bracket but possibly some natural regeneration on poorer sites that was not yet of commercial size; and the last category would include all other pine forest. In all cases, the forest stand encompassed by each training sample site should be uniform in respect to the criteria that were established for both the age/vigor category and the vegetation cover type in respect to predominant species. In the case of annual vegetation such as an agricultural crop for which all fields in a given region are likely to have been planted within a two to three week time span, the difference in age is not strongly correlated with vigor. Therefore, for agricultural crops, vigor is addressed in respect to whether a particular crop

is growing at the normal rate versus growing under a stress caused by insect/disease infestations, inadequate moisture, or lack of nutrients derived from the soil. If, at a particular time for which a classification is to be made, stress conditions are known to exist, training sample sites for a particular crop should be established both in fields in which plant growth is normal and in fields that are under some kind of stress. Because of likely differences in vigor, training sample sites for the same crop should be established for both irrigated and nonirrigated conditions, if both irrigation and dryland practices are intermingled for a given crop. Also, if there are other differences in practices such as a particular grass species or species mix being grazed in one case and grown for hay in another case, separate training sample sites should be established for each.

A span of several weeks in the planting period for a particular agricultural crop could cause variation in plant density which is the third principal variable in vegetative cover that may influence the reflected energy measured by the multispectral scanner. Plant density is an important factor because the multispectral scanner is taking a measurement for a 1.1 acre area. Consequently, if, as in the case of row crops as seen from above, there are both plants and exposed soil between the rows, the measurement involves an integration of energy as will be reflected

differently from plants and exposed soil. If a wide difference in planting dates for a particular crop indicates that a significant variation in density could exist, then training sample sites should be established for each density category. For practical purposes, such density categories should be fairly broad; e.g., 40% to 60%, 60% to 80%, and 80% to 100%. If the crop vegetation covers less than 40% of the surface, it is not likely that the specific crop could be identified. Vegetation density is also a factor in forest and brush vegetation, but it is recommended that the variation in density be addressed with two categories rather than three as recommended for agricultural crops. For example, training sample sites established to address density variation in pine forest may relate to criteria that cause them to be categorized as sparse (20% to 65% of the surface covered by tree crowns) or dense (65% to 100% of the surface covered by tree crowns). This example leads one to the fourth factor that must be considered in view of major variation in vegetation that influences reflected energy--that of the understory vegetation. In the previous example of a sparse pine forest, one could expect that in one condition there may be a native grass under the trees that would be visible from above in the gaps between the sparsely scattered trees, but that in-another condition the understory may be a brush species or species association rather than grass. In this example, it would be desirable to establish a training sample site

for each of the two conditions. In the case of Landsat data acquired during the winter season when deciduous forests are leafless it is also desirable to establish a training sample for dense, deciduous forest that has an evergreen understory species and another for dense, deciduous forest that has a deciduous understory species. Two training sample sites would also be appropriate for a leafless deciduous forest flooded with water versus a leafless deciduous forest without flooding.

Finally, the topography in respect to slope and aspect can be a factor in establishing criteria for the establishment of uniform, homogeneous training sample sites if there is pronounced topographical variation. In the case of pronounced slope, it is recommended that slope categories be established for 0 to 10% slopes, 10% to 30% slopes, 30% to 50% slopes, and 50% or greater slopes. Aspect is usually not considered in the training sample site criteria unless slopes are 30% or greater in which case aspect is categorized in respect to the four cardinal directions. In actuality, most steep slope conditions are in forested areas and are likely to be automatically categorized as to aspect in the course of applying criteria for defining the species or species association cover type. For example, in western United States a pronounced slope with a south aspect may support ponderosa pine; whereas a pronounced slope with a north aspect may support larch-Douglas fir.

In summary, the main variation in vegetated areas as it influences the measurement of reflected energy with a scanner is (1) the species or species association, (2) the vigor/age, (3) the density, (4) the ground cover or understory in certain forest conditions as described above, and (5) the slope and aspect in areas with pronounced topographical variation.

For land cover types that are not vegetated, other conditions must be considered during the establishment of training sample sites. During the spring season, most cultivated areas are in some stage of soil preparation. Consequently, the training sample site criteria are established in respect to the measurement of energy reflected from exposed soil rather than in view of the anticipated crop. The three main variables to consider in the establishment of training sample sites are the state of the surface, soil moisture, and soil type. As a minimum, training sample sites should be established to represent the extremes, should they exist, of these three variables and various combinations thereof. For example, the state of the surface should be considered in respect to a rough surface that may have resulted after plowing versus a smooth surface resulting after harrowing and/or planting; soil moisture in respect to dry conditions in some fields versus very wet or waterlogged conditions that may exist in other fields; and soil type in respect

to soil extremes such as light-colored, sandy soils in some fields versus dark-colored, clay soils in other fields. In the case of land used for cultivated crops, there may also be ground cover conditions other than green, growing vegetation or exposed soil that must be addressed at certain times of the year for which Landsat data are to be processed. One common condition is a stubble condition resulting from harvesting operations widespread in early fall. Even though there should be no significant difference in energy reflected from dead stalks and debris from different crops, one should establish training samples to separately represent various possible stubble conditions. For example, one should establish training sample sites to represent a stubble left after corn has been cut for silage in which there is a low volume of stalk material left and considerable bare soil exposed, separately from training samples representing stubble left from harvesting small grains.

Land cover types other than those that support vegetation--agricultural crops, pasture/grass, orchards, forest, brushland, marsh--can be addressed under the general heading of "inert materials." The category of inert materials would include beaches, sand bars, mud flats, rock outcropping, extractive areas (e.g., gravel pits), asphalt, concrete, etc.; in essence, those land cover types that are devoid of vegetation. Except for the topographical configuration, there is little variation within each of these land cover types; however, their basic characteristics may relate to different degrees

of reflectivity in the four MSS measured wavelengths. For example, concrete is highly reflective; whereas, asphalt has low reflectivity. Consequently, training sample sites should be established to represent each of these inert materials that may be present. In some cases these inert materials may not exist in pure form over an area large enough to serve as a training sample site (to be addressed in next section), and, therefore, a training sample site may be established to represent a particular complex. For example, a training sample site that contains a heterogeneous mixture including concrete streets, gravel parking lots, metal roofs, etc., would be appropriate in an area where these materials only exist in such a mixture. In the urban environment, some training sample sites may be termed "high density" to reflect a criteria requiring pure or mixed inert materials with no vegetation intermingled; whereas, others may be termed "low density" to reflect a criteria permitting up to 35% of the total surface to be covered by isolated patches of vegetation (no larger than 100 feet in maximum dimension) with the remainder encompassed by pure or mixed forms of inert materials. The former may typify large urban commercial centers or industrial sites, and the latter may typify suburban residential areas with scattered trees partially overlapping the streets and houses.

The first step in planning a ground truth information gathering operation for computer-implemented land cover classification with Landsat data consists of determining

and defining the major land cover categories that relate to the anticipated application(s). In this step the basic factors that influence reflected energy, as measured with the Landsat multispectral scanner, must be kept in mind so that the land cover categories are compatible with the data acquisition and processing technique. The second step consists of listing the variation that is anticipated in each land cover category so that a training sample site can be established to represent each source of variation. Again this listing should be derived in view of the basic principles elaborated in this paper and according to preestablished criteria similar to that used in the previous examples. Appendix A shows a typical outcome in respect to such a listing derived for Mississippi. It is, of course, important to understand that the listing of major land cover types must accompany a specific definition of each, and that the training sample site criteria must be predefined. Appendix B gives such definitions and criteria as examples relative to the terminology used in the appendix A listing. The reader will note that some vegetation types shown in the appendix A listing do not occur at all times during the year. Consequently, only part of these categories would be found during a ground truth operation conducted during a particular time of the year. Experience at ERL has shown that once the vegetation types occurring at a given time of the year have been brought together with variation in age, density, understory, and topography, there are typically around 30 to 50

land cover conditions found within a 115 mile by 115 mile area encompassed by a Landsat scene.

Size and Shape of Training Samples

There are two principal factors that relate to the size of the training sample site. One factor concerns the facility with which a training sample site can be located in the land surface image displayed on a cathode ray tube (TV-like screen) as the image is reconstructed from Landsat digital data. The other factor relates to the number of data cells (1.1 acre areas) for which the scanner takes a measurement of reflected energy that are needed in order to develop valid statistics from the energy measurements. Experience at ERL indicates that it is most desirable to establish training sample sites that are around 40 acres (16 hectares) in size. A training sample site of this size can usually be located in the land surface image displayed on the cathode ray tube (CRT) without difficulty; and will encompass around 30 data cells (1.1 acre areas), thereby providing a sample large enough to develop valid statistics. Towards the upper extreme, it is not recommended that training sample sites for natural vegetation be larger than about 160 acres (64 hectares), because it is likely that there would be too much difficulty in finding such a large site without violating the training sample site criteria for uniformity. Towards the lower extreme, the smaller the site the higher the probability that it will not be located in the Landsat data. Also, the

smaller the site the lower the efficiency in developing valid statistics--that is to say, two 20-acre sites with 15 data cells each could eventually be grouped in data processing to equal the 30 data cells encompassed by one 40-acre site, but this would require twice the effort in field work and data analysis. Consequently, it is not recommended that training sample sites smaller than 40 acres be established unless the particular land cover type in question does not exist except on areas smaller than 40 acres. In any event, it is recommended that 10 acres be the absolute minimum size for reasons of efficiency, statistical validity, and probability of locational accuracy in the Landsat data. In view of this recommended restriction, it can be seen that land cover types that only occur on areas smaller than 10 acres should be precluded from the list of land cover types that can be addressed with Landsat data. One should recognize, however, that once valid statistics are derived for a training sample and the spectral signature is developed for a particular land cover type within a Landsat scene, the classification will be performed in respect to each individual 1.1 acre data cell in the Landsat scene. (Provided that cloud-free conditions permit processing the four corresponding computer-compatible tapes as a data set.) Therefore, even though there may be some large areas in a 115 mile by 115 mile scene within which a particular land cover type only occurs in units between 1.1 and 10 acres in size, these units may be classified

accurately through use of training sample sites of adequate size found elsewhere in the scene.

The shape of training sample sites is not crucial; however, in some cases a square or rectangular-shaped training sample site is easier to locate in the Landsat data. As will be further elaborated in a diagram later in this report, locating sites in the CRT display of the land surface image is often facilitated by visually projecting lines from prominent, easily identified surface features to two or more of the sides of a square or rectangular site.

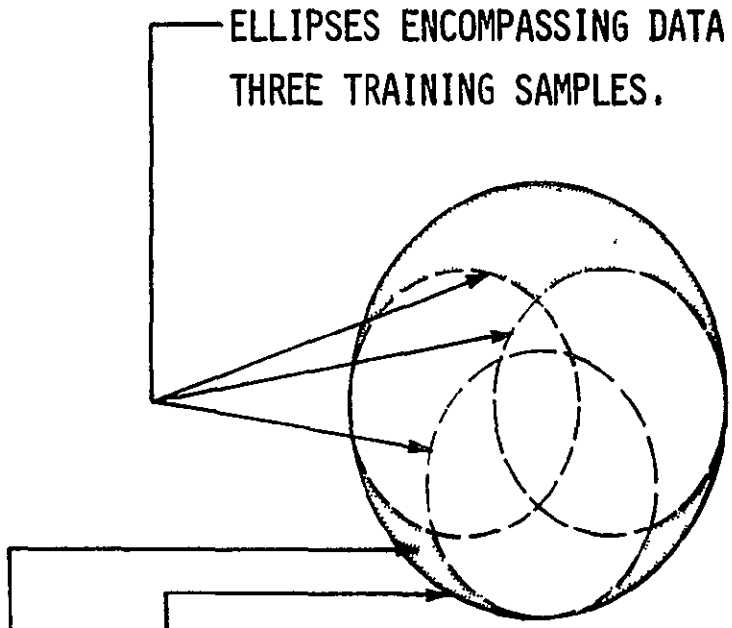
Number and Distribution of Training Samples

From a theoretical point of view, only one training sample is needed to develop a spectral signature and, subsequently, perform a computer-implemented classification of the land cover feature that the particular training sample site was established to represent, provided that the training sample statistically represents the land cover type to be classified. However, for several reasons, it is recommended that an attempt be made to establish at least three training sample sites for each land cover feature to be classified. First, it is possible that a training sample site could be lost either because its location cannot be established in the data or, infrequently, because its location coincides with scan line dropout--an electronic or transmission failure during which no measurements are recorded for all or part of the data cells on a particular

scan line. Secondly, it may be necessary to discard a training sample because the statistics, once derived from the Landsat data, indicate that it is not a uniform homogeneous land cover type from a spectral viewpoint. Such statistics may have resulted from human error during training sample site location either as delineated on field maps or located in the Landsat data, or it may be due to the basic nature of the particular land cover type. In any event, unless the problem can be corrected, the training sample is discarded. Consequently, if only one training sample site had been established, data processing would have been interrupted to redefine the boundary or by another field trip to establish a new site for the particular land cover condition. Thirdly, the analysis of the statistics is easier if the statistics from three or more training sample sites established to represent the same land cover condition can be compared. For example, if the mean and standard deviation, as calculated from the Landsat data, for one training sample is significantly different from that of the others, the analyst may discard that training sample on the basis that the remaining samples that coincide are a better representation of the particular land cover condition or he may carry it as a separate spectral subclass. On the other hand, if the analyst were dealing with only one training sample he would have no basis of comparison and would have to accept it as being representative; or, if he were dealing with two for which the statistics were

substantially different, he would not know which of the two was most representative. Finally, even though the statistics from three or more training samples established to represent the same land cover condition should not be substantially different, the means and standard deviations are usually not exactly the same. Consequently, the analyst may wish to group the three or more samples to create new statistics to develop a spectral signature that encompasses more variation in the particular land cover type. This theory is illustrated in diagram 1. In the example in diagram 1, the three dashed-line ellipses represent areas that encompass the measurements from all the data cells in each sample as they cluster around the means in the center of each ellipse. The solid-line ellipse constructed around the three dashed-line ellipses represents a hypothetical situation resulting from grouping the statistics of the three individual samples. Consequently, if the measurements for an unknown data cell fixed its location in the shaded area during classification, that data cell would be classified as pertaining to the particular land cover type; whereas, it would have been left uncategorized had each of the three training samples been carried as separate classes. In this hypothetical case, it can be seen that three grouped training samples would have resulted in a more accurate classification than one or two either held separately or grouped, if the grouped statistics more correctly estimated the true statistical population for the ground cover condition.

DIAGRAM 1



ORIGINAL PAGE IS
OF POOR QUALITY

NOTE: This diagram was meant to illustrate the concept rather than to imply that the solid-line ellipse resulting from the grouping of three training samples is always tangent to the individual sample boundaries. In fact, its position will be dependent on the confidence interval defined.

A ground truth information gathering operation is usually oriented to a particular Landsat scene that encompasses a 115 statute mile by 115 statute mile area (about 8-1/2 million acres) that relates to a set of four computer-compatible tapes (CCT's) that contain the digital data for computer-implemented land cover classification. The number of training sample sites needed within a particular Landsat scene varies with the number of land cover types to be classified within the scene and the variation within each land cover type. As discussed earlier in this paper and indicated in the listing in appendix A, there may be up to 11 major land cover categories in a state as large and varied as Mississippi, and it may be necessary to establish 80 or more training samples in order to address the variation in all conditions within these major land cover types during all seasons of the year. However, ERL experience indicates that within the area encompassed by a particular Landsat scene during a particular season for which ground truth information is being gathered, there are likely to be 8 to 10 major land cover types for which around 30 to 50 training sample sites must be established to address the various land cover conditions. Consequently, inasmuch as it is recommended, as previously discussed, that three or more training sample sites be established for each land cover condition, around 90 to 150 training sample sites may be established in the 8-1/2 million acre area encompassed by each Landsat scene. Using the upper extreme of this example and assuming that

each training sample site is 40 acres in size, one can calculate that 40 acres times 150 training sample sites amounts to less than one one-thousandths' of the 8-1/2 million acres encompassed by one Landsat scene (4 CCT's) being within training sample site boundaries.

As stated, a set of training sample sites would normally be established for the area encompassed by each Landsat scene because the four CCT's relating to a scene are usually processed as a set. However, in some situations, it may be possible, through an approach referred to as signature extension, to process more than the four CCT's from a scene as a set, and, thereby, reduce the number of training sample sites per scene. This possibility could arise in a situation where two or three cloud-free scenes (8 to 12 CCT's) of data were acquired on a particular pass under fairly uniform atmospheric conditions over the area of concern. This situation is most often encountered when the passage of a strong cold weather front precedes a Landsat pass by one or two days. However, it is recommended that ground truth information gathering activities be planned for the area encompassed by each scene, and that the concept of signature extension be considered only in respect to data processing efficiency after Landsat data have been acquired and assessed as to quality.

Training Sample Site Homogeneity/Uniformity

The fundamental requirement of the computer programs used to perform a land cover classification is that the

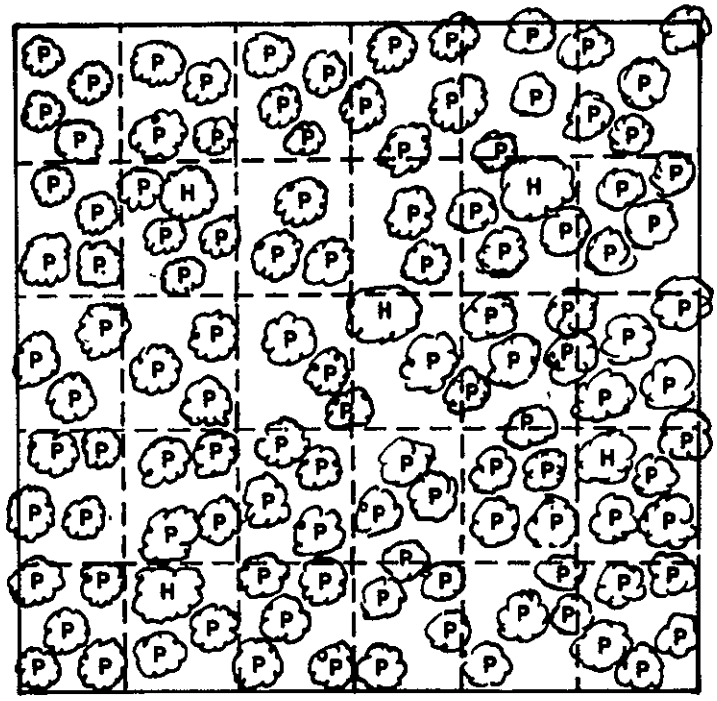
statistics derived from the multispectral scanner data that correspond to a training sample conform to a normal distribution. These statistical parameters are, then, used to establish an elliptically-shaped decision boundary which, in turn, is based on a normal distribution. If one allows statistics that do not reflect a normal distribution to define the decision boundary, the classification will be degraded. Therefore, it is necessary for a training sample site to be uniform and homogeneous in respect to the vegetation/land cover condition that it is selected to represent. The uniformity/homogeneity specification is made in respect to those vegetation/land cover variables that influence the reflected and/or radiant energy being measured by the multispectral scanner as elaborated previously. However, the concept is most easily addressed in diagram form.

Diagrams 2 and 3 show large squares with solid lines that are meant to represent training sample sites of about 40 acres in size. Within each large square, dashed lines are used to form small rectangular areas that are meant to represent the 1.1 acre cells for which the scanner takes a measurement.

Diagram 2 represents a training sample site established to represent a sparse (20% to 65% crown coverage) pine forest (90% or more pine) with a native grass ground cover apparent in the gaps between the trees. The circles with the

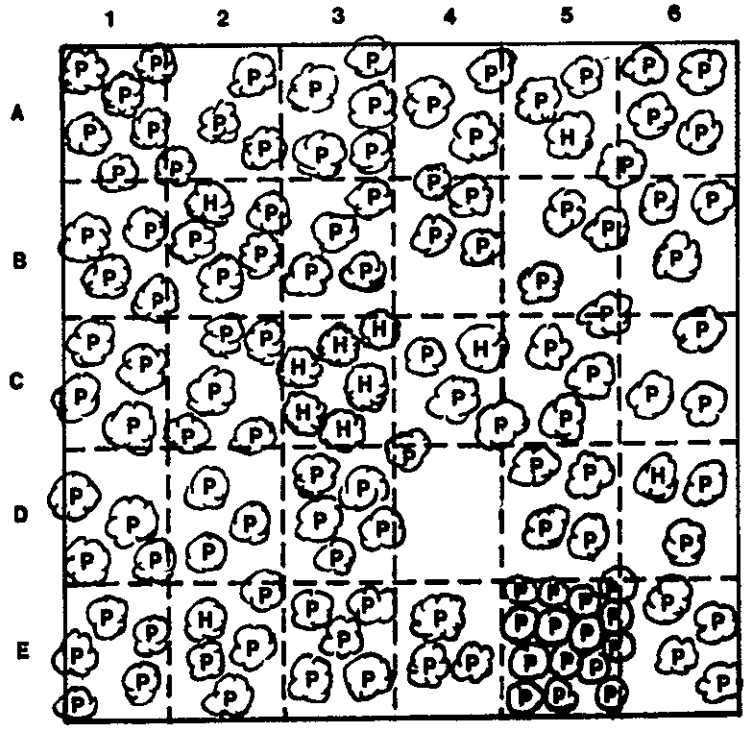
DIAGRAM 2

ADEQUATE



P - PINE
H - HARDWOOD

INADEQUATE



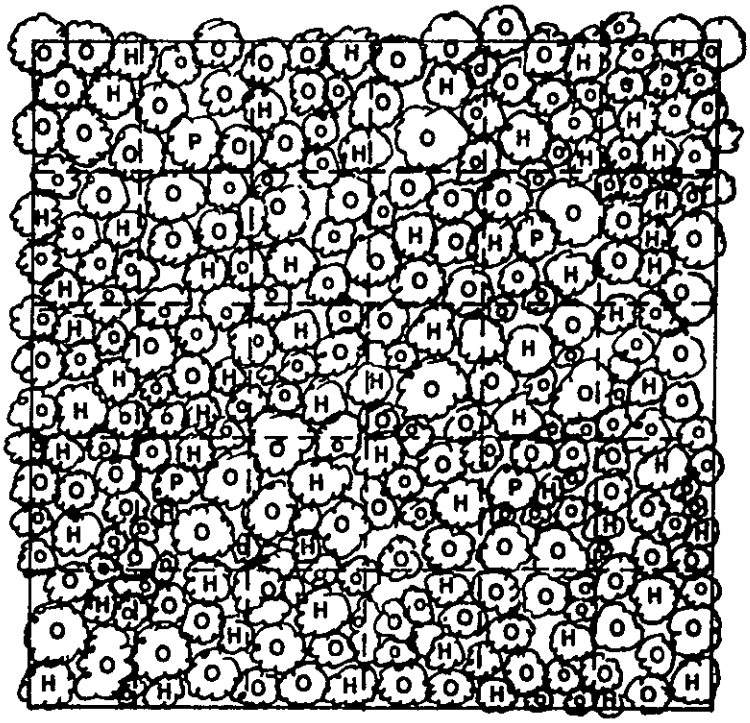
ORIGINAL PAGE IS
OF POOR QUALITY

letter P represent the areas covered by crowns of pine trees and the circles with the letter H represent the areas covered by crowns of hardwood trees as seen from above. The training sample site represented by the bottom square is not adequate as a uniform, homogeneous site for several reasons. In cell 3C there is a concentration of hardwood trees that encompasses the entire cell, in cell 4D the gap between trees is so large that the entire cell would fall on grass between the trees, and in cell 5E the density of the pine trees is such that it exceeds the criteria for a "sparse" condition. Conversely, the top square is adequate in respect to reflecting a uniform, homogeneous condition in that pine trees are scattered throughout the site in a manner that some fall in each cell without exceeding the criteria for a "sparse forest. Even though there are a few hardwood trees present they do not occur in concentrations and do not exceed 10% of the area covered by tree crowns, and there are no large gaps in the canopy.

Diagram 3 represents a training sample site established to represent a dense oak-hickory forest (90% or more oak-hickory). The training sample site represented by the bottom square is not adequate as a uniform, homogeneous site for two reasons. First, there is a concentration of pure oak in cells 2B and 3B; and, secondly, there is a concentration of pine in cell 4D. The top square is an adequate representation of a uniform, homogeneous condition in that both oak and hickory trees occur in each cell in roughly the same proportion; and, even

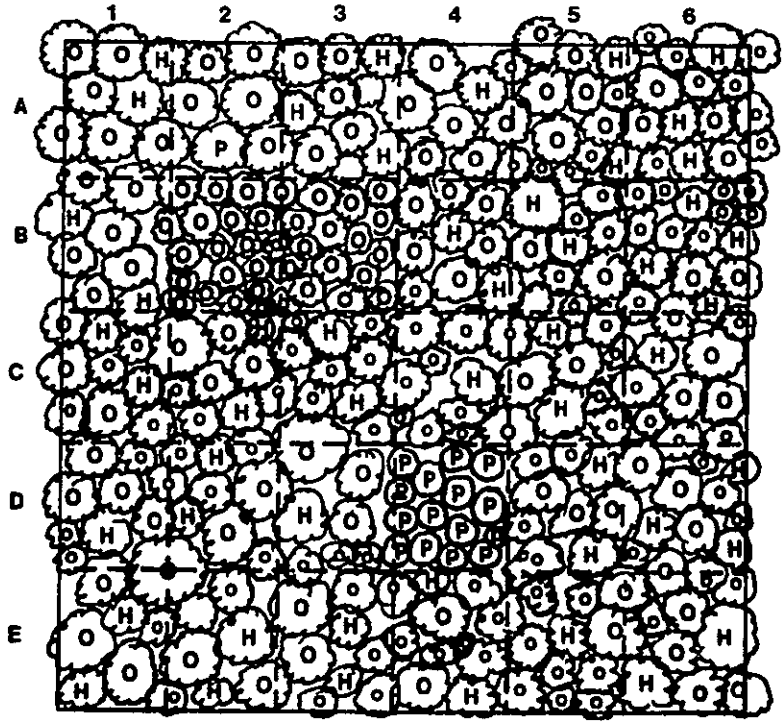
DIAGRAM 3

ADEQUATE



O - OAK
H - HICKORY
P - PINE

INADEQUATE



ORIGINAL PAGE IS
OF POOR QUALITY

though there are a few pine trees present, they are not concentrated and do not exceed 10% of the total area covered by tree crowns. With these diagram examples, it can be seen that the main characteristic for uniformity is that there be no condition within an area as large as 1.1 acres that differs substantially from the criteria that defines the land cover condition to be represented by the training sample site. Although the diagrams used forest vegetation as an example, the same criteria should be applied to other types of vegetation. For example, a marsh (non-forested wetlands) or a pasture grass characterized as a species association with two or more intermingled species should be such that none of the species occur singly over an area as large as a cell. It is also recommended that a species occupying less than 25% of the area not be included in the name of multi-species associations; and, therefore, would not be considered in applying uniformity criteria.

Agricultural crops are usually single species, but conditions of density and/or vigor may have a bearing on uniformity and homogeneity. For example, areas one acre or larger in size with bare soil due to germination failure or with differences in vigor due to uneven fertilization or poor nutrient availability should not be permitted in a training sample site established to represent an otherwise healthy crop.

Topographic features should also be uniform in respect to broad categories as suggested earlier in this paper. For

example, in mountainous terrain, a training sample site should not be established so that part is on a north aspect slope and part on a south aspect slope if slopes are greater than 30%, nor should the slope exceed the limits defined for a slope category; e.g., 30% to 50%.

III. PRE-FIELD IMPLEMENTATION PROCEDURES

The first step in implementing a ground truth information gathering operation consists of determining and defining the major land cover/vegetation categories, and the second step entails listing the variation in conditions that is anticipated in each major category as previously discussed.

The next step usually consists of using available aerial photography to preselect the required number of training sample sites according to predefined criteria for each land cover/vegetation condition that is to be represented by a training sample site. Although the preselection of training sample sites through air photo-interpretation is not essential, it can be used to gain efficiency in the ground truth information gathering operation.

Although any type of aerial photography that is available is adequate provided that it is not too old (within last five years under most conditions of land use change), ERL experience shows that preselection is most efficient with color infrared positive transparencies in

roll form at scales of 1:60,000 to 1:120,000 when interpreted under magnification. This efficiency is gained because large areas can be viewed on a single frame, the resolution at these scales is compatible with locating the training sample in Landsat data, and the logistical planning of field work is often facilitated when using one print that shows the road network for a large area. However, scales in the range of 1:15,000 to 1:30,000 can also be used with considerable efficiency and in some cases can have certain other advantages in respect to the kind of detail that can be photo interpreted.

In general, during a training sample site preselection process, the photo-interpreter does not strive to photo interpret all details on which ground truth information is desired. For example, he may find and delineate potential training sample sites that meet the criteria for pine forest, hardwood forest, marsh, and brushland, but stop short of identifying the particular species or species association. He may also delineate potential training sample sites for cultivated areas or grassland that appear to be uniform, leaving the ultimate categorization to the field team. In essence, the photo interpreter strives to add efficiency by preselecting training sample sites that meet the general criteria so that the field personnel (that may also include the photo interpreter) can go directly to these preselected sites as opposed to canvassing the entire area in search of adequate training sample sites.

In the end, even if the field personnel reject some of the preselected sites and reestablish substitute sites while in the field, the overall operation is usually less time-consuming than it would have been had sites not been pre-selected. However, depending on the type, scale, and season of acquisition of available aerial photography, the photo interpreter may deal with certain vegetation cover type variables more precisely than they can be dealt with on the ground. For example, color infrared positive transparencies acquired during the winter season at scales of 1:30,000 or larger can be used to determine density (crown closure) categories in pine forest and/or the degree of the overstory mix between pines and leafless hardwoods as precisely and with much less effort than can be determined on the ground. Broad slope and aspect categories can also be efficiently determined through stereo vision interpretation of forward overlapping photography.

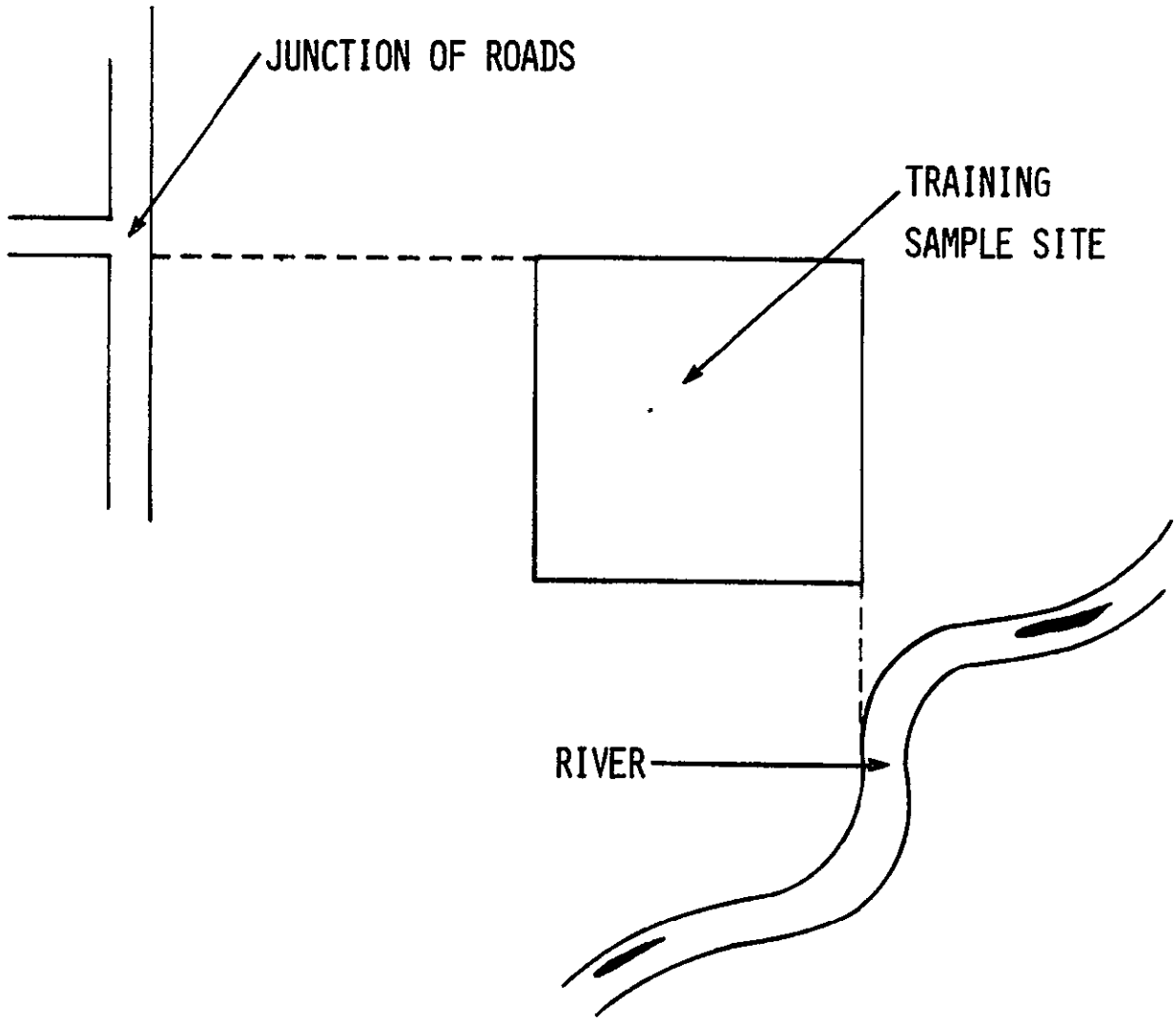
If the photo-interpreter is not familiar with general land cover/vegetation types within the area of concern, it is often helpful to review publications that give statistical information by county such as published by the U. S. Soil Conservation Service, the U. S. Forest Service, and the U. S. Crop Reporting Service.

As potential training sample sites are located through photo interpretation, the boundaries of these preselected sites are usually delineated on transparent material overlaid

on the original film for office records, and on prints made for use in the field. If the original photography was color infrared or color, black and white prints are usually made for field use; and, if the original scale is smaller than one inch to the mile (1:63,360), a print enlarged to about 1:63,360 is made to facilitate use in the field. As potential training sample sites are preselected and delineated, the photo-interpreter writes a unique four to six digit letter/number identifier on the print and overlay adjacent to the delineation of the site. This unique letter/number identifier, to be explained later in this section, is used both for cross-reference to ground truth forms and for identifying the site during computer processing.

Another means of increasing overall efficiency by preselecting training sample sites through air photo-interpretation is introduced by having the photo-interpreter delineate potential sites in a manner that their location is referenced to prominent surface features that are easily found in the field and detectable in the image display of Landsat data. This concept is illustrated in diagram 4 in which the potential training sample site was so delineated that one side can be located by visually projecting from a road junction and another side can be located by visually projecting from a bend in a river. If such linear features are 50 feet or wider they can almost always be used, subsequently, to easily locate the training sample site in the

DIAGRAM 4



Landsat image display. Other means of using this concept even more effectively is to delineate potential training sample sites so that one or more sides are adjacent and parallel to straight-line interfaces between two different land cover types (e.g., forest and cropland); or, in the case of cropland and grassland, delineate training sample sites in fields with one or more sides parallel and immediately adjacent to prominent roads. It is also often possible to project lines from the centers of two or more prominent non-linear features such as small water bodies or built-up areas as references to the sides of potential training sample sites. ERL experience has shown that, if due attention is given to this concept during preselection of sites and/or in field establishment of sites, very few, if any training sample sites are "lost" because their locations cannot be ascertained in the image display of Landsat data; thereby, gaining both time and ultimately, cost efficiency in the overall operation.

During preselection of potential training samples, the photo-interpreter should also observe the road network and, whenever possible, locate potential training sample sites so as to facilitate access and take best advantage of road networks. Attention of this sort will reduce field work time by cutting down walking time and backtracking vehicle routes.

A final means of gaining efficiency consists of establishing potential training sample sites in concentrated

groups distributed throughout the area encompassed by a particular Landsat scene of interest. This can be accomplished by having the photo interpreter begin by selecting 8 to 12 air photos, depending on scale, from the entire set of photography available for the area encompassed by a particular Landsat scene. This selection can be made so that each photo encompasses a variety of land cover types or in a manner that each photo focuses on a particular land cover type depending on whether field teams are organized along multi-disciplinary or disciplinary lines. In either case, it is desirable that each air photo covering an area of concentrated training sample sites fall completely within the area covered by one of the four CCT's relating to the particular Landsat scene to preclude the establishment of training sample sites at the abutment of tapes. Also, if field personnel are organized relative to political or management units (e.g., a county forester), it is also desirable that all or most of the air photos fall within that particular unit. If the photo-interpreter does not encounter a sufficient number of potential training sample sites that meet the predetermined criteria for each land cover condition with the original selection of air photos, he can select additional air photos so as to optimize the chances of finding sites relating to land cover conditions lacking after the first iteration. The net effect of delineating potential training sample sites in concentrated groups distributed throughout the Landsat scene should be

the reduction of travel time between sites during field operations; and, subsequently, the reduction of time required to locate training sample sites in the image display of Landsat data.

At the present time, aerial photography that has been taken within the last five years and is of a type and scale suitable for training sample site preselection is available for most of the United States. If not already in the possession of agencies planning Landsat ground truth information gathering operations, the existence and coverage of aerial photography acquired through various Federal programs can be ascertained through and purchased from the EROS Data Center at Sioux Falls, S.D., operated by the U. S. Geological Survey. Landsat coverage for a particular area defined by latitude and longitude can be verified through and purchased from the EROS Data Center. If recent aerial photography is not available for a particular area of interest, it may be cost-effective to acquire a limited amount of new aerial photography especially in forest or marsh areas with poor accessibility. If the acquisition of new aerial photography is carefully planned, it should be possible to acquire sufficient aerial photography to gain cost-efficiency in ground truth information gathering operations as discussed in this report by covering no more than 2% of the area with aerial photography (e.g., 36 frames or 12 sets of 9-inch format stereo triplets at 1:24,000 scale per Landsat scene). If the aerial photography is to be used for ground

truth information gathering activities for specialized (e.g., pine versus hardwood stratifications for forest inventory) rather than composite classifications, limited color infrared photography during the winter season at a time close to a cloud-free Landsat pass could completely eliminate the need and cost of field activities.

Even if aerial photography is not used for preselection of potential training sample sites, it is desirable to provide air photo prints to field personnel to be used as a map base on which to delineate the field-established training sample sites. In the case of the ASVT project to which this report relates, preselection of potential training sample sites through photo interpretation was accomplished within seven counties for which specific applications were demonstrated, but ground truth information gathering operations were conducted without preselection in the remaining 75 counties of the State of Mississippi. However, the field personnel were provided with either air photo prints or photo base maps with broad land cover type delineations. The original photography was 1:120,000-scale color infrared, but the prints were reproduced in black and white to reduce cost and enlarged to 1:60,000 to facilitate use in the field. A print of one photo (10 mile by 10 mile effective area) was provided for each of 65 counties, and a township size (6 mile by 6 mile) photo base map with land cover types delineated was provided for 10 counties. The net effect was to concentrate established

training sample sites in either "10 mile by 10 mile" or "6 mile by 6 mile" areas within each county. Inasmuch as the training sample sites were established by field personnel assigned to each respective county (e.g., county extension agents and county foresters) or a management unit within a county (e.g., state park or game management unit), all field personnel were very familiar with their respective area of responsibility. In some cases, the field personnel were so familiar with their areas that, after orientation of the air photo print, they could delineate satisfactory training sample sites, each representing some specific vegetation/land cover condition in their area and fill out corresponding ground truth forms without leaving their office. It is mainly in this situation, in which field personnel are very familiar with a localized area, that training sample sites can be established efficiently without preselection through photo-interpretation.

Prior to field implementation of a ground truth information gathering operation, ground truth packages should be assembled for field personnel. In the case of this ASVT project, the packages prepared for disciplinary personnel located in each county consisted of (1) an air photo or photo-based land cover map as described previously in this report, (2) a county map that shows the outline of the area encompassed by the air photo, (3) various blank ground truth forms (to be discussed later in this report), (4) an instruction sheet, and (5) a

sheet showing letter symbols for each vegetation/land cover condition that could be characterized by a four-digit combination of letters. For example, a pine forest that was old (50 years +) and sparse was characterized as PFOS. Instructions called for field personnel to write the appropriate letter symbols adjacent to each training sample site delineated on the air photo print. Instructions also called for a unique two-digit number to be added to the four letter characters as each training sample site was established and delineated on the air photo print. For example, 09 added to PFOS would mean that PFOS09 was the ninth training sample established. This six-digit identifier was also brought over to the ground truth form (which contained additional information) corresponding to a particular training sample site as a cross reference. In the case of the seven counties for which preselection of training sample sites through photo interpretation was conducted, the air photos taken to the field had one to four letter symbols, depending on the degree to which photo-interpretation of land cover conditions was possible, together with a unique two-digit number recorded adjacent to the delineation of the potential training sample site.

In the situations when preselection of potential training samples through photo-interpretation is conducted and when the field team is not very familiar with the local area, it is recommended that the appropriate location of each preselected site be plotted with an X on a small-scale map such

as a 1:250,000 topographic quad map. This map can then be used for logistical planning in order to assign field teams to specific areas outlined on the map in view of site locations, the road network, and lodging facilities.

IV. FIELD IMPLEMENTATION PROCEDURES

Organization of Field Personnel

There are various possible ways to organize a field team for gathering ground truth information. This report will focus only on what are considered to be the two basic options. One basic option is to organize the effort around field personnel employed by those agencies that are the anticipated users of the land cover classification. This option was followed in the course of this ASVT project resulting in the incorporation of field personnel as summarized in table 1. As is implied by table 1, this option involves an organized effort in which each individual is responsible for establishing training sample sites within his local area for his area of specialty. For example, a county forester would only establish training sample sites to represent the various forest vegetation conditions within the county to which he is assigned. With this form of organization, each individual involved would use only a fraction of his time for the establishment and visitation of training sample sites, and most work could be conducted in the course of carrying-out routine activities as opposed to a separate effort.

TABLE 1 -- FIELD PERSONNEL BY TYPE OF GROUND TRUTH

CATEGORY	COORDINATING AGENCY	FIELD PERSONNEL	GEOGRAPHIC AREA
CROPS, PASTURE, ORCHARDS	MS. COOP. EXTENSION SERVICE	82 COUNTY AGENTS	STATEWIDE
COASTAL WETLANDS	MS. MARINE RESOURCES COUNCIL	3 GULF RESEARCH LAB.	COASTAL ZONE
OTHER NATURAL VEGETATION	MS. FORESTRY COMMISSION	63 COUNTY FORESTERS	STATEWIDE
	MS. GAME & FISH COMMISSION	8 DISTRICT BIOLOGISTS	20 GAME MGM'T AREAS
	MS. PARK COMMISSION	15 PARK SUPERINTENDENTS	15 STATE PARKS
URBAN & BUILT-UP	MS. R&D CENTER/ECO. DEVELOP. DISTRICTS	10 ECONOMIC & DEVELOPMENT DISTRICTS	URBAN AREAS
EXTRACTIVE	MS. GEOLOGICAL SURVEY	2 JACKSON OFFICE TECHNICAL STAFF	STATEWIDE

ORIGINAL PAGE IS
OF POOR QUALITY

The other basic option for carrying-out a ground truth information gathering operation consists of employing a team representing several disciplines (e.g., a forester, an agronomist, and a botanist) whose primary responsibility would be to gather ground truth information. The possibility of this option is envisioned in a situation in which it would be feasible to make such a team part of the staff of a remotely sensed data processing center. As will be explained later in this report, it is thought that three disciplinary personnel could furnish ground truth information for a state as large as Mississippi utilizing 80% of their time, leaving 20% of their time for performing certain steps during data processing and interfacing with disciplinary personnel from user agencies.

There are advantages and disadvantages associated with each of these two basic options for organizing a ground truth information gathering operation. In respect to the first option discussed, one advantage involves capitalizing on local field personnel's detailed knowledge of local vegetation/land cover conditions and road systems. In addition, it is thought that local field personnel, who eventually become users of the classification products, could make better use of the products having become familiar with the manner that such products were produced by having been personally involved in the process. Finally, the utilization of local field personnel substantially reduces the funds needed for travel and per diem expenses. The principal disadvantage of utilizing a large number of local field personnel is that it involves a substantial coordination

effort. It is thought that the magnitude of the coordination effort for a state as large as Mississippi and with a comparable number of operating agencies would be such that a designated coordinator would expend at least 25% of his time in coordinating activities related to ground truth gathering. In addition, the effort required to orient a large number of local field personnel in ground truthing procedures is substantial in respect to both time and cost.

Conversely, the second basic option discussed has the principle advantages that the little coordination required could be accomplished by the team itself, and the team could be formed by personnel already trained in the use of remotely sensed data. In addition the team can be utilized to give continuity to the total operation from ground truth gathering through the location of training sample sites in the Landsat data and in performing analysis activities during data processing more effectively than local field personnel could be utilized to attain this continuity. The main disadvantage of a small, centralized team is the "cost per training sample site established" may be higher due to additional travel/per diem costs and some lost field time caused by the lack of familiarity with local road systems and conditions.

Distribution of Responsibility

When the ground truth is to be gathered by a large number of local field personnel, it is most important to have a well conceived plan to distribute responsibility. However, because

such a plan must take account of the exact organization of field personnel to be involved, it is impossible to provide any more than general guidelines.

Ideally, training sample sites that are established to relate to a particular Landsat scene should be distributed so that some fall on each of the four computer-compatible tapes for the particular scene. Diagram 5 shows the nominal Landsat coverage of the two principal tracks over Mississippi with the solid-lines areas representing Landsat scenes and the dashed-lines showing areas relating to the four tapes within each scene. For simplification, overlap between tracks is not shown but the amount of overlap can be estimated in reference to the distance between the dashed-line in the center of each scene and either of the dashed-lines on each side of the center line. Because there is some shifting in the Landsat coverage from pass to pass and because areas covered by individual tapes may not encompass all of a land unit (e.g., a county) to which field personnel relate, it is not practical to assign responsibility for areas corresponding to each tape. Especially, if field personnel are assigned to counties or management units within a county (e.g., a game management unit or state park), a practical and simple manner of assigning responsibility is to request that each field person establish a given number of training sample sites in each land cover condition within his respective county. This approach was followed in this ASVT project in a manner that county field personnel were requested to establish one training sample site for each vegetation/land

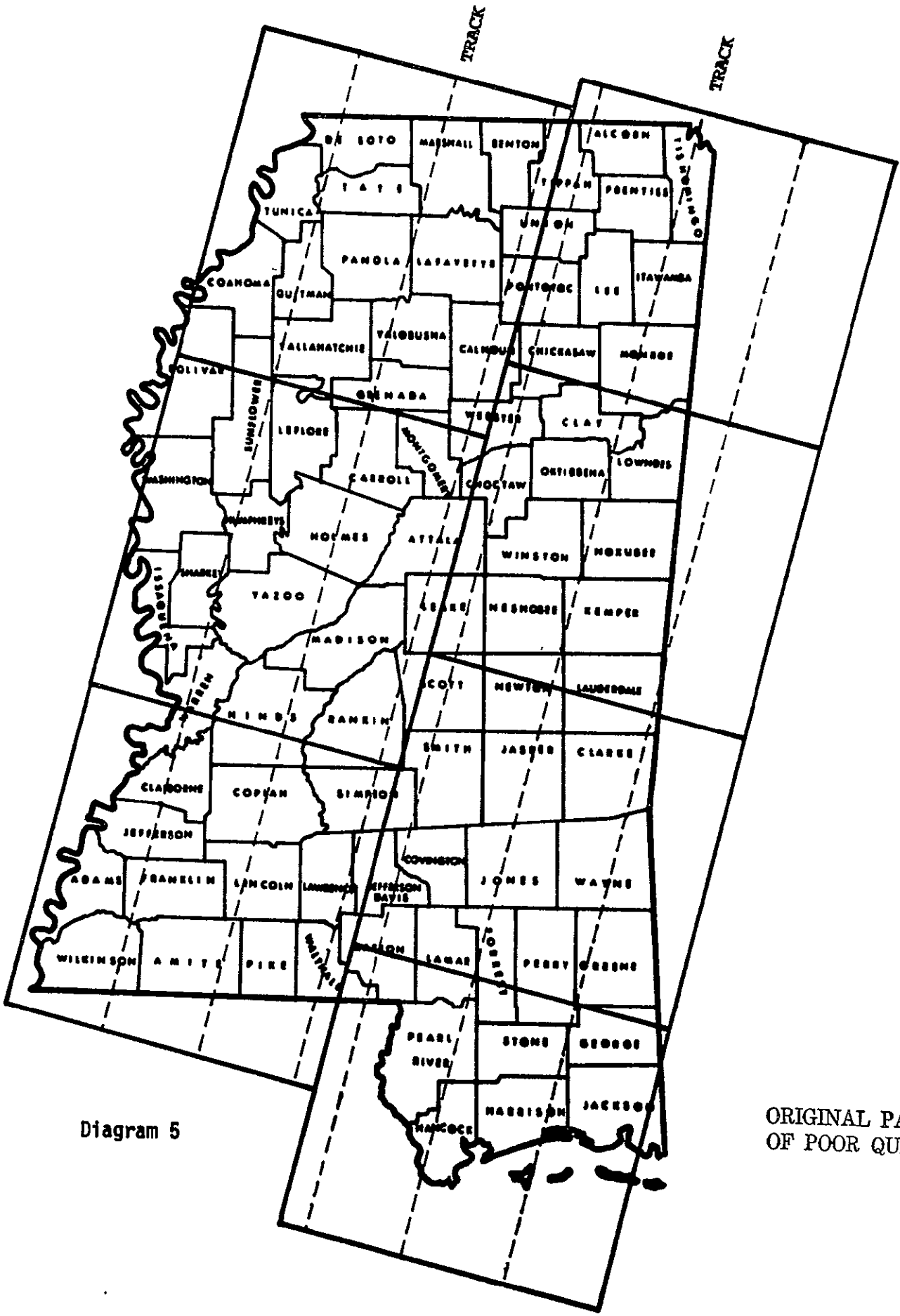


Diagram 5

ORIGINAL PAGE IS
OF POOR QUALITY

cover condition occurring within the area encompassed by an air photo or map selected for each county. Inasmuch as field personnel were organized along disciplinary lines, the effect of the rule was that county foresters established one training sample site for each forest vegetation condition, county extension agents established one for each cropland and pasture condition, etc. The effect of supplying field personnel with one air photo or township map within the respective county was to get some concentration of ground truth within each county and, thereby, save time during the location of training sample sites in the Landsat data at the same time that a distribution of ground truth throughout the Landsat scene was attained. The size of counties in Mississippi is such that, on the average, there are twelve counties within each Landsat scene; therefore, with a rule of one training sample site per land cover condition per county, twelve training sample sites for each land cover condition are theoretically possible. However, because all land cover conditions did not occur within the area covered by the air photo or township map selected for each county, the actual outcome varied from three to eight per land cover conditions per Landsat scene.

In other situations where the average size of counties is substantially larger or smaller than the average Mississippi county, the guideline used for the ASVT project could be adjusted accordingly. For example, in the case of smaller counties, if there were from 20 to 24 counties per Landsat scene, one may select one-half the counties for ground truthing

and still use the simple instruction of one training sample site for each land cover condition. The set of four CCT's for a given Landsat scene are usually processed as one data set; however, because of cloud problems, it may be necessary to use one tape from one scene acquired on a given date and three tapes from another scene acquired on a different date. Consequently, it is desirable to select counties for ground truthing in a manner that established training sample sites are likely to occur in the areas encompassed by each tape in the nominal Landsat scene. In addition, if there are different physiognomic areas within a particular Landsat scene, counties should be selected to be somewhat proportional to the area encompassed by each physiognomic unit. For example, in Mississippi, a Landsat scene may encompass both an alluvial plains agricultural area and uplands area with mixed land use--in which case counties for ground truthing would be selected to represent each of these two physiognomic units roughly in proportion to the extent of each.

Timing of Ground Truth Gathering

Although ground truth information gathering can take place during any time of the year, it is desirable to restrict field activities to be within a prime time defined for each season; thereby, avoiding transitions in respect to seasonal change and/or agricultural land use. For example, a transition between winter and spring may capture forest vegetation in a state that deciduous trees are neither leafless or fully leafed-out, and

agricultural fields in a state when some were stubble from the previous crop, some are being plowed, and some are being planted.

Although the commencement or termination of the phenomena associated with a given season may fluctuate somewhat year to year due to weather patterns, the prime time for ground truth gathering for each season is generally considered to be as follows in Mississippi:

Season	Prime Time
Winter	January 15 - February 28
Spring	April 15 - May 30
Summer	July 1 - August 15
Fall	October 15 - November 30

Inasmuch as the accuracy with which various land cover types can be classified varies between seasons, ground truth for specialized classifications should be conducted during the prime time for that season. For example, ground truth for a classification of coastal marsh vegetation should be conducted during prime time for the summer, the best time to separate pine forest from other vegetation is during winter, for agriculture crops during summer, etc. If a good composite classification was desired with one set of data, a ground truth gathering operation during the spring season would be most appropriate. However, the best possible vegetation/land cover data base could be built up with a classification during each season, followed by subsequent update as need (Joyce,A.,1974).

Another consideration as to the timing of ground truth information gathering concerns the means by which field activity is instigated. There are several possible alternatives. One alternative is to make a "go/no go" decision based on observations of cloud cover at the time of each satellite overpass during the prime time. In this situation a "go" decision would be made for the first cloud-free or relatively cloud-free (95% to 100%) pass after which field personnel would be immediately notified to gather ground truth information within 10 to 15 days. This means of instigating field activity assures that ground truth will be close to the date for which Landsat data are acquired for processing, but requires a high degree of coordination between the weather observers, the decision maker, and the field personnel. In addition, it limits the amount of time that field personnel have to perform their work. Another alternative is to preselect a scheduled Landsat pass date during prime time, and instruct field personnel to gather ground truth within plus or minus a given number of days (e.g., 10 to 15) from that date. This alternative is easier to implement and gives the field personnel more flexibility in planning and conducting their activities to fit their own schedule, but has the disadvantage that the cloud condition may not be acceptable on the preselected overpass date. Of course, this does not preclude using ground truth acquired in this manner to process data from another pass closest to the preselected date. A third alternative is to instigate field activities to occur within a defined six week prime

time period irrespective of satellite overpass dates. This instruction offers the greatest flexibility for field personnel to schedule and conduct their work and is the easiest to implement, but increases the chance that some training samples will have to be discarded during data processing because land cover conditions changed between ground truthing and Landsat data acquisition. The latter approach was followed in this ASVT project with the end result involving a discard rate of only 3% of all training samples due to apparent change in conditions.

Ground Truth Information Forms

Section III of this report discussed the need to develop "ground truth information forms" to be used by field personnel It is recommended that separate forms be developed for each major land cover category or associated categories as opposed to developing a single form for all land cover categories. For example, one form may be prepared for forest and brush vegetation, another for pasture and crops, another for urban areas, etc. Separate forms of this nature allow disciplinary field personnel to deal only with forms pertinent to their responsibility, but also can be developed in a simpler format and reduce the total bulk of paperwork to be handled. Examples of forms used for this ASVT project are shown in appendix C. Another option in the development of forms concerns a "check-off" versus a "fill-in-the-blank" approach. It is thought that a "check-off" approach is most desirable because it not only saves

time but is much easier to use under field conditions. For most information, the "check-off" approach is easy to develop and use; however, for some information the "fill-in-the-blank" approach may be necessary. In the case of natural vegetation, unless the person developing the form is aware of all possible species associations in the area of concern, it is better to use a "fill-in" approach so as not to preclude obtaining ground truth on some. However, even when the "check-off" approach is used, field personnel should be instructed to establish training sample sites for any land cover condition encountered that meets training sample site criteria, even though not indicated on the form.

Orientation of Field Personnel

Pages 36 and 37 of this report list the contents of a package of materials assembled for each field person. Although this package contained an instruction sheet (see appendix D for an example) that was meant to exclude the necessity for verbal explanation as to field procedures, it is desirable to hold orientation meetings with designated field personnel to deliver the package, review all details of its contents, outline areas of responsibility, discuss timing of ground truth gathering, etc. In the case of this ASVT project, orientation meetings were held at various locations throughout the state usually in the district offices of each agency involved. A total of 15 orientation meetings were held with from 8 to 18 field personnel participating in each meeting. Each meeting averaged about three hours with

the first hour used to explain the basics of satellite data acquisition and processing and the last two hours used to review the ground truth package contents, explain procedures, areas of responsibility, etc. However, if time and travel funds permit, it is more desirable to assemble field personnel for orientation meetings at the location of the data processing equipment, and provide them with a full day of orientation. This would allow a system demonstration including the mechanics of locating training sample sites in the CRT display of the Landsat data image. Experience at ERL shows that such a demonstration gives field personnel a better feel for how the location of training sample sites can be located through reference to other features in the image, and visually emphasizes the need to establish training sample sites that are uniform and homogeneous in respect to the land cover condition each is to represent.

Field Work

The essence of field work, associated with a ground truth information gathering activity, is to verify or establish the location of each training sample site that is uniform and homogeneous in respect to the land cover condition that it is established to represent, and to fill out a ground truth information form for each site. If potential training sample sites were preselected through photo-interpretation, the field person simply locates the delineated area on the ground, verifies that the area delineated is uniform and homogeneous,

and, if so, makes the necessary observations to fill in the ground truth information form. If the field person finds the delineated area to be inadequate in respect to uniformity criteria, he may discard it and locate and delineate another area. In other cases he may find it suitable to erase part or all of the delineated boundary and delineate a new boundary shifted slightly from the original delineation. In some cases the field person may find that the photo-interpreter delineated a uniform and homogeneous vegetation condition but errored in the interpretation and named it something other than what it actually was. For example, the photo interpreter using photography acquired during the winter season when hardwood trees were leafless may have been misguided by an evergreen understory component (e.g., holly or wax myrtle) apparent through the leafless overstory and called it pine forest. In this case, the field person could simply change the letter/ number identifying symbol recorded by the interpreter to the correct symbol and fill out the form accordingly. As a basis for filling out the ground truth information form, once that an adequate training sample site is located on the ground, the field person may take various approaches. On one extreme, in the case of a training sample site delineating an agricultural crop in a 40-acre field bounded by roads on two or more sides, he may make most observations from a vehicle stopping only to make two or three spot checks by walking into the field. On another extreme, in the case of natural forest vegetation, he may use pacing and a hand compass to keep his bearings as he

follows some pattern to assure adequate coverage (as suggested in appendix E) along which he stops occasionally to make visual observations. As the location and delineation of training sample sites and corresponding ground truth information form completion proceeds, it is extremely important that the letter/number identifier symbol (as described in section III of this report) be recorded both on the ground truth form and on the air photo print or map adjacent to the delineation of the corresponding training sample site. In addition, it is very helpful if field personnel staple all ground truth forms to the air photo or map on which training sample sites corresponding to those forms are delineated. As explained previously in this report, it is desirable to delineate training sample sites on recent aerial photo prints or photo-based maps; however, if such are not available, training sample sites can be delineated on 7-1/2 minute series (1:24,000) topo maps, or, in the absence of those, 15 minute series (1:62,500) topo maps if such maps are not so old that they are grossly out-of-date. However, it is recommended that maps at scales smaller than 1:63,360 (1 inch to the mile) should not be used for training sample site delineation.

If potential training sample sites are not preselected through photo-interpretation, the operation usually starts with local field personnel delineating some sites on air photos or maps in the office based on their knowledge or office records of the area. Visitation of these sites as well as delineation and visitation of additional sites can usually be

performed in the course of the local field personnel's routine work. However, if the time period indicated for the ground truth gathering activity is short, a special effort may be required. After sites are located, the work proceeds in the same manner as described for sites preselected through photo-interpretation.

In the case when personnel are employed exclusively for ground truth gathering, it is most efficient if they function as a team by meeting at the end of each day to keep a master list of training sample sites established and plan the next day's activity. In this manner, preselected sites that may have been rejected or lost because of access problems may be substituted for by another team member.

The involvement of field personnel in producing a land cover classification with Landsat data may end with delivery of air photo prints with delineated training sample sites and corresponding ground truth information forms. However, it is desirable that field personnel also assist in the location of training sample sites in the Landsat data. ERL experience has shown that assistance from field personnel can save time both through more rapid location of sites in the display of the Landsat image on the CRT and in catching possible recording errors.

Once that training sample sites have been established for the first land cover classification, ground truth information

for additional classifications can be obtained with substantially less effort. Except in the case of agricultural land on which use changes from year to year, it is only necessary to ascertain that no drastic change has occurred since the first ground truth effort; consequently, ground truth forms can be greatly simplified. An example of a ground truth form prepared for revisits of training sample sites established for forest vegetation is shown in appendix F.

Time Required and Cost

A tally of time required to make observations within a training sample site, delineate the site on an air photo or map, and fill out the ground truth information form showed the following distribution for this project:

Time Required	No. of Sites
5 to 15 minutes	93
15 to 30 minutes	130
30 to 60 minutes	117
1 to 2 hours	37
2 to 3 hours	8
Over 3 hours	5

There was a noticeable difference in time required for training sample sites for different land cover conditions. On an average, crop and pasture sites required 24 minutes per site; whereas, forest and brushland sites required 43 minutes. It was not possible to keep account of travel time and expense (vehicle

operating and depreciation costs) because most sites were established during field personnel's routine work. However, ERL experience shows that a field person can be expected to establish and provide ground truth on an average of six training sample sites per day when travel time between sites within the county is included. Consequently, for programmatic purposes, it can be estimated that 25 man-days or 200 man-hours would be required to address up to 150 training samples that may be established for one Landsat scene with 13,260 square miles.

Using a rate of $10.50 per hour to cover all costs (salaries, overhead, and operating costs) it can be calculated, based on this assumption, that ground truthing would cost $0.16 per square mile (200 hours x $10.50 ÷ 13,260 square miles) for the first classification. This calculation is compatible with past cost calculations at ERL which, although derived in a research environment rather than an operational environment, indicate a range from $0.15 per square mile for the easiest ground truth gathering (recent air photos, field personnel familiar with the area, and easy access/terrain) to $0.31 per square mile for difficult ground truth gathering (no air photos, field personnel not familiar with the area, and difficult access/terrain). (See Joyce, A., 1975.) Actually, it is not realistic to prorate all costs for the first ground truthing effort against the first classification because revisiting the established training sample sites for subsequent classifications requires far less effort.

In certain situations, such as when dealing with large, inaccessible marsh or wetlands areas, ground truth gathering can be more costly. It is in these conditions that selected coverage with new aerial photography (if not available) and/or the use of helicopters should be considered. However, even after cost comparisons are made with costs assuming access by boat, a higher cost for use of helicopters may be considered an adequate trade-off in view of the time required for a limited number of personnel to use boat transportation.

Although ground truth for this project was acquired by local field personnel, it was mentioned earlier in this report that an option would be to utilize a ground truth team that would work almost exclusively for this purpose. Such an effort for a state as large as Mississippi may involve a breakout of work activity as follows:

Days	Activity
150	Photo Interpretation and Pre-field Preparation
180	Field Work and Travel Within Counties
150	Post-field Records and Location of Sites in Landsat Data
48	Travel from Central Location to Counties
528	Total

Allocating 220 work days per year for each of three persons (e.g., an agronomist, a forester, and a botanist) equals 660 days leaving about 20% of their time for interface with users on other activities (e.g., digitizing other information). However, it may be most desirable to utilize local field

personnel for the first complete ground truthing effort and a full-time team of three for revisits and update for subsequent classifications and specialized classifications. In this manner, the first ground truthing operation could be accomplished rapidly, and the field personnel, who eventually become users of the land cover classification, become familiar with the manner in which the classification is derived from Landsat-acquired data.

V. CONCLUDING COMMENTS

This report addressed ground truth information gathering procedures relative to performing a land cover classification using Landsat multispectral scanner digital data and the "supervised" approach to computer-implemented data processing. One difference between "supervised" and "cluster analysis" approaches is that, in the former, ground truthing takes place first so that the computer can be "trained" to recognize a land cover condition elsewhere; whereas, in the latter, classification takes place first and, then, a ground truthing operation is launched to determine the land cover condition that corresponds to each resulting class. Modified approaches employing unaided training sample selection and supervised classification are also in use. However, inasmuch as the same basic data are utilized and the basic principles involved in the measurement of reflected and/or emitted energy are the same, this report should have relevance to ground truthing activities irrespective of the approach to data processing.

It is not likely that new techniques will cause drastic changes to the procedures outlined in this report in the near future. However, techniques that are currently being developed and/or tested may cause some slight changes in ground truth gathering procedures. When raw data are registered to a given map projection so as to permit the development of techniques to allow automated location of training sample sites in the Landsat data, it will be necessary to determine the map coordinates (e.g., UTM) that define the location of each training sample site. If techniques currently under development and testing at ERL to define categories of mixed vegetation through spatial analysis of classified data is successful, this may preclude the need to establish training sample sites for some mixed vegetation categories (e.g., an oak/pine mix). Also, when techniques to merge land cover information from seasonal classifications into a master composite classification are perfected, it may be desirable to conduct ground truth information gathering activities during each season of the year in a manner that each seasonal activity is specialized so as to encompass only those land cover categories that can be most accurately classified with Landsat data acquired during the respective season (see Joyce, A., 1974).

This report was written primarily to help field personnel understand the principles and procedures involved in ground truth information gathering. However, it is thought that it would be beneficial for field personnel to familiarize themselves with aspects of data processing and analysis

(see Butera, K., 1976; Joyce, A. and Griffin, R., 1976 and Whitley, S., 1975). Such familiarization would not only enhance their understanding of ground truth information gathering, but would also give them a better understanding of both the advantages and limitations of using land cover classifications derived from Landsat data.

LITERATURE CITED

1. Butera, M. K., 1976; A Technique Development for the Determination of Louisiana Marsh Salinity Zones from Vegetation Mapped by Multispectral Scanner Data: A Comparison of Results From Satellite and Aircraft Data. ERL Report #161, NASA Earth Resources Laboratory, Slidell, LA 70458.

2. Jones, W. C., 1974; Implementation of an Advanced Table Look-Up Classifier for Pattern Recognition. Proceedings of the Ninth International Symposium on Remote Sensing of Environment, Vol II, Environmental Research Institute of Michigan, Ann Arbor, Michigan.

3. Joyce, A. T., 1974; Computer Implemented Land Use Classification with ERTS Digital Data During Three Seasons. ERL Report #135, NASA Earth Resources Laboratory, Slidell, LA 70458.

4. Joyce, A. T. and Derbonne, J. D., 1975; A Computer Implemented Land Use Classification Technique Applied with ERTS Digital Data Acquired Over Southern Missouri. ERL Report #143, NASA Earth Resources Laboratory, Slidell, LA 70458.

5. Joyce, A. T. and Griffin, R. H., 1976; The Use of Landsat Digital Data and Computer Implemented Techniques for an Agricultural Application. ERL Report #159, NASA Earth Resources Laboratory, Slidell, LA 70458.

6. Pendleton, T. W., 1976; Digital Overlaying of the Universal Transverse Mercator Grid with Landsat Data Derived Products. ERL Report #160, NASA Earth Resources Laboratory, Slidell, LA 70458.

7. Whitley, S. L., 1975; A Procedure for Automated Land Use Mapping Using Remotely Sensed Multispectral Scanner Data. NASA TR 4-434, National Technical Information Service, Springfield, VA 22151.

8. Whitley, S. L., 1976; Low-Cost Data Analysis Systems for Processing Multispectral Scanner Data. NASA TR R-467, National Technical Information Service, Springfield, VA 22151.

APPENDIX A

Land Cover Conditions in Mississippi

MISSISSIPPI TRAINING SAMPLE MASTER FILE CATEGORY LIST

CROPLAND
- CS = Soybeans
- CC = Cotton
- CN = Corn
- CR = Rice
- CM = Wheat
- CZ = Field Peas
- CA = Potatoes Sweet
- CB = Grain Sorghum
- CP = Peanuts
- CQ = Cumcumbers
- CW = Watermelons
- CH = Peppers
- CG = Forage
- CF = Crop Fallow
- CE = Exposed Soil, Cleared
- CI = Plowed
- CJ = Disked
- CK = Harrowed
- CO = Stubble
- CT = Oats

PASTURE AND HAYLAND
- IB = Bermuda
- IC = Bahia
- ID = Dallas
- IF = Fescue
- IA = Alfalfa
- IE = Combination
- IT = Temporary (e.g. Ryegrass)
- IO = Other

NATURAL GRASSLAND
- NF = Native Field Grass
- NW = Native Woodland Grass

ORCHARDS
- OC = Citrus
- ON = Pecans
- OP = Peaches

ORIGINAL PAGE IS
OF POOR QUALITY

FOREST LAND
- PS = Longleaf-Slash
- PL = Loblolly-Shortleaf
- PP = Pine Plantation
- FP = Oak-Pine Mix
- HE = Leafless Hardwood W/Evergreen Understory
- HD = Leafless Hardwood W/Deciduous Understory
- HO = Oak-Gum-Cypress
- HH = Oak-Hickory
- HW = Willow
- HC = Cypress-Tupelo
- HB = Maple-Beech-Birch
- HA = Elm-Ash-Cottonwood
- HM = Hardwood Mixed
- HP = Hardwood Plantation

BRUSHLAND
- BH = Brush Deciduous
- BE = Brush Evergreen
- BM = Brush Mixed
- BD = Brush Debris (e.g. recent clearcut)

MARSHLAND
- BA = Baccharis Halimifolia
- BB = Baccharis Halimifolia/Golden Rod
- DS = Distichlis Spicata/ Scirpus
- JA = Juncus Roemerianus/Spartina Cyanosuroides
- JB = Juncus Roemerianus/Baccharis Halimifolia
- JC = Juncus Roemerianus/Distichles Spicata
- JD = Juncus Roemerianus/Spartina Alterniflora
- JS = Juncus Roemerianus/Spartina Patens
- ME = Cyperus/Eleocharis Cellolsa
- SC = Spartina Cyanosuroides/Scripus
- SJ = Spartina Alterniflora/Juncus Roemerianus
- TY = Typha

EXTRACTIVE
- EG = Gravel Pit
- ES = Sand Pit
- EZ = Sand/Gravel Pit
- EQ = Quarry/Limestone
- EM = Strip Mine, coal
- EC = Clay Extraction

INERT MATERIALS
- IS = Sand Beach or Bar
- IM = Mud Flat
- IH = Hard Surface (asphalt, concrete)
- IG = Building
- IB = Barren or rock outcrops

URBAN BUILTUP
- UH = High Density
- UL = Low Density

WATER
- WR = River
- WD = Deep Lake, Reservoir
- WS = Shallow Lake, Reservoir
- WC = Catfish Pond
- WO = Other

APPENDIX B

Definitions and Criteria

DEFINITION OF MAJOR LAND COVER/VEGETATION TYPES

CROPLAND - A specified unit area which is usually planted to an agronomic crop or grass on an annual basis after soil preparation.

PASTURE/GRASSLAND - Specified unit area of which 90% or more of the surface covered with foilage is covered with foliage of grasses, generally used for grazing or hayland on other than an annual basis.

FORESTLAND - Specified unit area of which 10% or more of the surface area is covered with foliage of trees.

PINE FOREST - Forest in which 66 2/3% or more of the area covered with foliage of trees is covered by foliage of evergreen trees as seen from above.

HARDWOOD FOREST - Forest in which 66 2/3% or more of the area covered with foilage of trees is covered by foliage of deciduous trees as seen from above.

MIXED PINE/HARDWOOD - Forest that does not meet the above criteria for evergreen or deciduous forest.

BRUSHLAND - Specified unit area of which 90% or more of the surface area covered with foliage is covered with foliage of multi-stemmed, perennial shrub species.

FORESTED WETLANDS - Forested areas that are seasonally flooded for prolonged periods (usually three months or more) of the year and/or flooded due to diurnal tidal action directly or indirectly through water backup.

MARSHLAND - Specified unit area that is frequently inundated for prolonged periods and contains plant species typical of "nonforested wetlands" areas over 90% or more of its surface.

ORIGINAL PAGE IS
OF POOR QUALITY

SPECIES ASSOCIATION - A vegetation type in which two or more plant species grow intermingled with the foliage of each species covering at least 25% of the surface area as seen from above.

SPARSE CROWN CLOSURE - Forested area in which 10% to 65% of the surface is covered by crowns (foliage and branches) of overstory trees when in leafed condition.

DENSE CROWN CLOSURE - Forested area in which 65% to 100% of the surface is covered by crowns (foliage and branches) of overstory trees when in leafed condition.

APPENDIX C

Ground Truth Forms

GROUND TRUTH DATA FOR FOREST, BRUSH, AND ORCHARDS

TAKEN BY:_____ DATE:_____

TRAINING SAMPLE IDENTIFIER_____ MAP OR AIR PHOTO INDEX #_____

ESTIMATED FIELD SIZE:_____ft X _____ft. or _____ACRES

LOCATION_____
County ¼ ¼ Section Township Range

KIND OF VEGETATION (Check One)
( ) Natural Forest
( ) Forest Plantation
( ) Brush Vegetation

IF NATURAL FOREST, INDICATE:

(1) Major forest type (check one)

( ) Maple-Beech-Birch	( ) Elm-Ash-Cottonwood	( ) Aspen-Birch
( ) Oak-Hickory	( ) Loblolly-Shortleaf	( ) Oak-Pine
( ) Oak-Gum-Cypress	( ) Longleaf-Slash	( ) Mixed Hardwood

(2) Overstory Crown Closure

( ) Dense (65% to 100%) ( ) Sparse (10% to 65%)

(3) Overstory species composition (to nearest 25%)

Species	%

(4) Understory species compostion (to nearest 25%)

Species	%

(5) Average age class of upper canopy trees (check one)

( ) Less than 20 years	( ) 50 to 100 years
( ) 20 to 50 years	( ) over 100 years

(6) Average height class of upper canopy trees (check one)

( ) Less than 20 feet	( ) 50 to 100 feet
( ) 20 to 50 feet	( ) over 100 feet

(FOREST, BRUSH, ORCHARDS CONTINUED)

(7) Slope (Check One)

( ) 0% to 10%
( ) 10% to 30%
( ) 30% to 50%
( ) 50% or more

(8) Predominant Aspect (Check One)

( ) North ( ) South ( ) East ( ) West

If Forested Wetlands are flooded at time of observation, indicate depth of water:

( ) less than 1' ( ) 2' to 4' ( ) greater than 4'

or if not flooded at time of observation indicate:

( ) appears subject to flooding by water backup due to tidal action.

( ) appears to have been flooded for a prolonged period prior to observation.

If Forest Plantation or Orchard, indicate:

Species_____ Average Age_____

Spacing_____ Average Height_____

Row Direction_____

If Brushland, indicate species composition to nearest 25%:

(1) Species	%

(2) Vegetation Density:

( ) Sparse, 10% to 65% of surface covered.

( ) Dense, 65% to 100% of surface covered.

(3) If sparse density, ground level is:

( ) Grass

( ) Exposed earth

( ) Other _____

GROUND TRUTH DATA FOR CROPS AND PASTURE

TAKEN BY _____ DATE _____

TRAINING SAMPLE # _____ MAP OR AIR PHOTO INDEX # _____

ESTIMATED FIELD SIZE: _____ ft X _____ ft. or _____ ACRES

LOCATION _____
County 1/4 1/4 Section Township Range

GENERAL CONDITION(1) _____

DESCRIPTION (if not crop or pasture) _____

_____

CROP OR PASTURE SPECIES(2) _____ VARIETY (if known) _____

PLANTING TECHNIQUE(3) _____ PLANT HEIGHT (to closest ft) _____

ROW WIDTH _____ PHYSIOLOGICAL STATE(4) _____

ROW DIRECTION _____ VISUAL ASPECT(5) _____

PERCENT GROUND COVER ( ) 0% to 20% ( ) 40% to 60% ( ) 80% to 100%
( )20% to 40% ( ) 60% to 80%

WEED INFESTATION (species & %, if greater than 20%) _____

DISEASE INFESTATION (kind & %, if greater than 20%) _____

INSECT INFESTATION (kind & %, if greater than 20%) _____

SOIL CONDITION(6) _____

SOIL MOISTURE(7) _____

SOIL TYPE(8) (if available) _____

_____

OTHER COMMENTS (if needed) _____

_____

(1) e.g. crop, pasture, stubble, plowed, fallow.
(2) e.g. soybean, bahia grass, etc.
(3) e.g. row, skip row, drilled, broadcast.
(4) e.g. flowering, heading, mature, etc.
(5) e.g. chlorotic, wilted, etc.
(6) e.g. freshly cultivated, rough, smooth, etc.
(7) e.g. moist, dry, waterlogged, etc.
(8) series, texture, color, slope, etc.

GROUND TRUTH DATA
Extractive Land Uses

OBSERVATIONS MADE BY _____ DATE _____

IDENTIFIER NO.* _____ Approx. Size _____ X _____ (feet) or _____ acres.

COUNTY _____

LOCATION (if known) _____
Township Range Section Quarter Forty

ACTIVITY TYPE

( ) Sand pit
( ) Gravel pit
( ) Stone, dimension
( ) Stone, crushed
( ) Lime
( ) Cement
( ) Clay
( ) Chert & Tripoli
( ) Lignite
( ) Heavy mineral
( ) Other _____

_____

Is area ( ) in-production or ( ) abandoned?

If abandoned, is area ( ) barren or ( ) revegetated?

Is the area likely to contain impounded water during all or a significant part of year ( ) yes ( ) no?

How much time did it take to make observations and fill out this form (min. and/or hours) _____.

*Observations should only be made on extractive areas that are at least 600 feet by 600 feet, or approximately 10 acres. Once such an area is located, its location should be delineated on an aerial photo or map sheet with colored pen or pencil, and an identifier cross-reference number should be recorded on the aerial photo or map beside the delineated area and on the ground truth data form.

GROUND TRUTH DATA FORM FOR URBAN AREAS[1]

Training Sample ID No._____

Collected by:_____ Date:_____

High Density Urban ( )[2]
Low Density Urban ( )[3]

If High Density Urban - Predominantly Concrete ( )
Predominantly Asphalt ( )
Predominantly Other ( ) _____e.g., metal roof
Inert Material Complex ( )

Comments:[4] _____

_____

If Low Density Urban
Main type of inert Material - Roof tops ( )
Concrete ( )
Asphalt ( )
Other ( ) _____

Main type of vegetation -
Grass (lawns) ( )
Pine trees ( )
Hardwood trees ( )
Mixed pine/hardwood ( )
Mixed grass/trees ( )

Comments:[4] _____

_____

(1) An urban area training sample should be 1000 ft. by 1000 ft. or larger; however, if homogeneous areas of such dimensions cannot be located, areas of 500 feet by 500 feet or larger (approx. a city block) are acceptable.

(2) High Density Urban is defined as an area essentially devoid of vegetation; but with up to 10% covered with vegetation in small scattered parcels whose largest dimension is generally less than 100 feet.

(3) Low Density Urban is defined as an area within which inert materials (roof tops, concrete, asphalt) are predominant; but with up to 45% of the surface covered with vegetation, including overtopping trees, occurring in small, scattered parcels with the maximum dimension of each parcel no greater than 200 feet.

(4) Appropriate comments include identification of scenario, e.g., airport runway, industrial complex, downtown commercial area, etc.; height of buildings, e.g., one or two story, three to five story, 6 or more stories; pitch of roofs, e.g., flat, moderate angle, steep angle; or any other information pertinent to measurements made with overhead remote sensors.

GROUND TRUTH FORM FOR MARSH VEGETATION

1. Sample number_____.

2. Date: _____.

3. Time: _____.

4. Vegetation type:

   (1) pure stand (monotypic)_____.

      (a) species:_____.

   (2) intermixed (less than 6 vascular species present) _____.

      (a) dominant species:_____, _____.

   (3) intermixed (more than 6 vascular species present) _____.

      (a) dominant species:_____, _____.

      (NOTE: If a species comprises less than 5% of vegetation do not regard as major or dominant component.)

5. Homogeneity:

   (1) sub-elements (defined)

      (a) vegetation differences (clumps, patches, zones) _____.

      (b) barren areas _____.

      (c) open water _____.

      (d) sparse vegetation/barren _____.

      (e) sparse vegetation/water _____.

      (f) other (describe) _____.

   (sub-elements (size)

      (a) less than 10 feet _____.

      (b) more than 10, but less than 20 _____.

      (c) more than 20, but less than 40 _____.

      (d) more than 40, but less than 60 _____.

(3) distribution (of sub-elements in study area).

(a) evenly _____.

(b) center _____.

(c) peripheral _____.

(4) density (of vegetation as % of surface area).

(a) dense > 90 _____.

(b) intermediate < 70 _____.

(c) sparse < 50 _____.

6. Height of plants (stands).

(1) approximate height of major units:

(a) species _____, height _____.

(b) species _____, height _____.

(c) species _____, height _____.

(2) approximate height of minor units:

(a) species _____, height _____.

(b) species _____, height _____.

7. Status of vegetation:

(1) approximate (%) of dead-standing material.

(a) major units (species) _____.

(b) minor units (species) _____.

8. Stage of growth:

(1) major units

(a) dormancy (winter-no leaves) _____.

(b) dormancy (winter-leaves dead-standing) _____.

(c) seedlings _____.

(d) immature _____.

(e) mature _____.

(f) anthesis _____.

(g) vigor _____.

(1) excellent _____.

(2) fair _____

(3) poor _____.

9. Surface of substratum:

(1) covered by algae _____.

(2) covered by small vascular plants _____.

(3) covered by detritus _____.

(4) barren _____.

(5) substrate type

(a) mud _____.

(b) sand _____.

(c) sandy/mud _____.

10. Water level.

(1) standing on surface of marsh _____.

(a) covered by tidal water _____.

(b) covered by river overflow _____.

(c) combination of both (a & b) above _____.

(d) permanent or semi-permanent _____.

(2) Depth of water on marsh surface _____.

11. Comments:

APPENDIX D

Instruction Sheet

PROCEDURE FOR ESTABLISHING "TRAINING SAMPLE AREAS" AND DOCUMENTING "GROUND TRUTH" FOR AGRONOMIC CROPS AND PASTURES

STEP #1 - Locate one typical field for each of the different crops and pastures that occur within the geographic area covered by the aerial photo (or photomap) provided. Each such field will be referred to as a "training sample area." (Note: a 10-acre field is the minimum size suitable for a training sample, but a larger field, 40 acres to 160 acres, is desirable.

STEP #2 - Outline each training sample area located in Step #1 with pen or pencil, assign a reference number to each (starting with the number one), and record the reference number on the aerial photo (or photo map) along side of each outlined field (training sample area).*

STEP #3 - For each training area outlined and referenced on the aerial photo (or photomaps) in Step #2, fill out one "Ground Truth Data" form. Information on the form that is not readily available or not applicable can be so indicated in the appropriate blank. Record the index number of map or airphoto print on which the training sample is located in the upper right hand corner of the form.

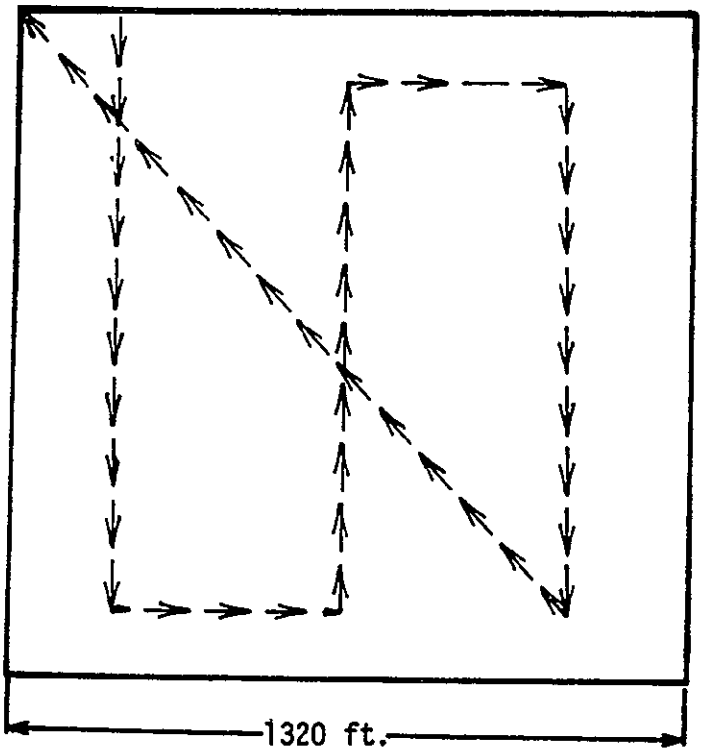
STEP #4 - Return all materials to project coordinator as soon as Steps #1 thru #3 are accomplished. This can take place between July and August; however, the earlier the better.

*If scale is 1:62,500 (air photo), a 40-acre field is roughly ¼" X ¼" on the photo; if scale is 1:24,000 (township map), a 40-acre field is roughly 2/3" X 2/3" on map.

APPENDIX E

Forest Uniformity Verification Pattern

Suggested Coverage Pattern
for ground verification of
uniformity for forest vegetation
training sample site of around
40 acres in size.



APPENDIX F

Revisit Form

GROUND TRUTH DATA FORM

SITE IDENTIFIER CODE (6-7 digit code recorded on air photo or land use map)_____

COUNTY:_____

OBSERVATIONS MADE BY:_____ DATE:_____

Has the vegetation within the training sample area delineated on the air photo or land use map been altered during the last year? ☐ yes ☐ no

1. yes, what was the cause?

- ☐ logging
- ☐ land clearing
- ☐ fire
- ☐ heavy insect or disease mortality
- ☐ other (indicate) _____

In which month did the alteration occur (if known)?_____

How much time did it take you to make observations and fill out this form (min. and/or hours)?

_____.

ORIGINAL PAGE IS
OF POOR QUALITY